

TITLE

STACKED GATE FLASH MEMORY DEVICE AND METHOD OF FABRICATING

THE SAME

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to a semiconductor device and method of fabricating the same. More particularly, it relates to a stacked gate flash memory device that can achieve high memory cell capacity.

Description of the Related Art

Complementary metal oxide semiconductor (CMOS) memory is generally categorized into two groups: random access memory (RAM) and read only memory (ROM). RAM is a volatile memory, wherein the stored data is erased when power is turned off. On the contrary, turning off power does not affect the stored data in a ROM.

In the past few years, market share of ROM has been continuously expanding, and the type attracting the most attention has been flash memory. The fact that a single memory cell is electrically programmable and multiple memory cell blocks are electrically erasable allows flexible and convenient application, superior to electrically programmable read only memory (EPROM), electrically erasable programmable read only memory (EEPROM) and programmable read only memory (PROM). Furthermore, fabricating flash memory is cost effective. Having the above advantages, flash memory has been widely applied in consumer electronic products, such as digital cameras, digital video cameras,

mobile phones, notebooks, personal stereos and personal digital assistants (PDA).

Since portability of these electrical consumer products is strongly prioritized by consumers, the size of the products must be minimal. As a result, capacity of flash memory must increase, and functions must be maximized while size thereof is continuously minimized. Having an increased amount of access data, capacity of memory cells has been enhanced from 4 to 256 MB, and even 1G byte will become the market trend in the near future.

Hence, there is a need for a flash memory device with high memory cell capacity.

SUMMARY OF THE INVENTION

Accordingly, an object of the invention is to provide a stacked gate flash memory device that can achieve high integration of memory cells thereof.

Another object of the invention is to provide a method of fabricating a stacked gate flash memory device, wherein the size of memory cells thereof can be reduced and the coupling ratio of the control gate to the floating gate can be also increased.

Moreover, the invention provides a method of fabricating a stacked gate flash memory device, wherein the driving currents thereof can be increased without increasing the surface size thereof.

Thus, a stacked gate flash memory cell in accordance with the present invention comprises a substrate having a trench therein. A conductive layer is disposed on the bottom of the trench. A pair of source regions are each respectively disposed in the substrate adjacent one sidewall

of the trench and electrically connected by the conductive layer. A source isolation layer is disposed on the conductive layer. A pair of tunnel oxide layers are each respectively disposed on one sidewall of the trench and the source isolation layer. A U-shaped floating gate is disposed on the source isolation layer, and connects the tunnel oxide layers thereby. A pair of control gate spacers are each respectively disposed on each vertical portion of the U-shaped floating gate, substantially having the same width as the vertical portions. A U-shaped inter-gate dielectric layer is disposed on the U-shaped floating gate and the control gate spacers. A control gate is disposed in the U-shaped inter-gate dielectric and a drain region is disposed in the substrate adjacent to the trench.

Furthermore, the method of fabricating the stacked gate flash memory cell in accordance with the present invention comprises providing a substrate, forming a plurality of parallel long trenches along a first direction in the substrate, forming a conductive layer and a pair of source regions on the bottom of each long trench, wherein the source regions are respectively disposed in the substrate adjacent to two sidewalls of each long trench and electrically connected by the conductive layer therein, forming a source isolation layer on each conductive layer, forming a tunnel oxide on two sidewalls of each long trench, forming a U-shaped floating gate on each source isolation layer, contacting the tunneling oxide layers, forming a pair of control gate spacers respectively disposed on the vertical portion of the U-shaped floating gate, substantially having the same width as the vertical portions, forming an U-shaped inter-gate dielectric layer on

each U-shaped floating gate and the control gate spacers, forming a control gate in each U-shaped inter-gate dielectric layer, forming a plurality of parallel shallow trench isolation (STI) regions along a second direction, defining a plurality of cell trenches and forming a drain region in the substrate adjacent to each cell trench.

In the present invention, the trench-type stacked gate flash memory device disposed in cell trenches within a substrate can achieve higher integration of memory cell capacity than that in the Prior Art.

In addition, most of the fabricating processes in the invention are self-aligned and additional lithography processes and number of masks for the whole fabricating process can be reduced. The complexity of fabrication is reduced and can be easily achieved. The higher coupling ratio by the memory cells also provides a lower operating voltage thereof.

Furthermore, most patterns of the masks for fabricating the stacked gate flash memory device are rectangular and can be easily fabricated. The costs of mask fabrication can be reduced and resolution limitations by the photolithography tools can be reduced.

A detailed description is given in the following embodiments with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention can be more fully understood by reading the subsequent detailed description and examples with references made to the accompanying drawings, wherein:

Fig. 1 is schematic top view of stacked gate flash memory devices of the invention;

Fig. 2a is a schematic top view of corresponding cross-sections in Fig. 3a and Fig. 4a for one embodiment of the invention;

Fig. 2b is a schematic top view of corresponding cross-sections in Fig. 3j and Fig. 4j for one embodiment of the invention;

Fig. 3a~3l are cross-sections of the fabricating process along the A-A' phantom line in Fig. 1; and

Fig. 4a~4l are cross-sections of the fabricating process along the B-B' phantom line in Fig. 1;

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a stacked gate flash memory device that meets the demand for increased capacity of memory cells. In Fig.1, a top view of the stacked gate flash memory device in accordance with the present invention is shown. Structure of individual memory cells is disposed in the cell trenches (referring to trench 250') along the A~A' phantom line, between two shallow isolation trench (referring to STI) regions along the B~B' phantom line.

Fig. 3a~3l and Fig. 4a~4l respectively illustrate the cross-sections of a fabricating process along lines A-A' and lines B~B' according to an embodiment of the present invention. Moreover, Fig. 2a and 2b also illustrate corresponding top views of the fabricating process.

First, Fig. 3l illustrates a cross-section of the stacked gate flash memory cells in accordance with the present invention. A memory cell comprises a trench 250' disposed in the substrate 200, and a conductive layer 206 disposed on the bottom of the trench 250'. A pair of source regions S are each respectively disposed in the substrate

200 adjacent to sidewalls of the trench 250'. A source isolation layer 207 is disposed on the conductive layer 206. A pair of tunnel oxide layers 208 are each respectively disposed on one sidewall of the trench 250' and the source isolation layer 207. A U-shaped floating gate 209 is disposed on the source isolation layer 207 and contacts the tunnel oxide layers 208 thereby. A pair of control gate spacers 211 are each respectively disposed on each vertical portion of the U-shaped floating gate 209. A U-shaped inter-gate dielectric layer 212 is disposed on the U-shaped floating gate 209 and the control gate spacers 211. A control gate 213 is disposed in the U-shaped inter-gate dielectric 212, and a drain region D is disposed in the substrate 200 adjacent to the trench 250'.

In Fig. 3a and 4a, a semiconductor substrate 200, for example a P-type silicon substrate, is provided. Next, a pad oxide layer 201 and a mask layer 202 are sequentially formed on the substrate 200 and then sequentially defined by subsequent lithography and etching, forming a plurality of long trenches 250 with a depth about 30000Å to 70000Å in the substrate 200. The long trenches 250 are parallel to each other along a first direction. This top view is shown in Fig. 2a and the depth of a long trench 250 affects the channel length of each stacked gate flash memory cell and the resistance of a source line.

Next, in Fig. 3b and 4b, a conformable bottom insulating layer 203 is deposited on mask layer 202 and in the long trench 250. The material of the insulating layer 203, for example, can be silicon dioxide. Then materials of a protective layer 204 are deposited on the bottom insulating layer 203 and fill the long trench 250. Material

of protective layer 204, for example, can be photoresist. The protective layer 204 is then etched back and recessed to a depth H to the surface of the substrate 200. The depth H is about 2500Å to 3500Å. Then the bottom insulating layer 203 exposed by the residue protective layer 204 is removed. Thus, a bottom insulating layer 203 having a same depth H from the surface of the substrate 200 as that of the protective layer 204 is left in the long trench 250.

In Fig. 3c and 4c, the residual protective layer 204 in the long trench 250 is then removed by a proper solvent (not shown). Next, a source material layer 205 is formed by sequential deposition, etch-back, and recession of impurities comprising doped materials such as N-type dopant doped polysilicon, for example, and preferably phosphorous (P) doped or arsenic (As) doped polysilicon. Thickness of the source material layer 205 is between 500Å and 1000Å. Next, a thermal annealing process, for example a furnace annealing process (not shown), is performed to drive out dopants (such as arsenic or phosphorous) from the source material layer 205 into the substrate 200 adjacent to the long trench 250. Thus, a pair of source regions S are each respectively formed in the substrate 200.

In Fig. 3d and 4d, the conductive materials of the source material layer 205 are further recessed to a depth H' from the surface of the substrate 200. Thus, a conductive layer 206 with a thickness H' is left in the long trench 250 and connects the adjacent source regions S by the conductive properties of the residual dopants (such as arsenic or phosphorous) therein. Moreover, a depth difference ΔH between the bottom insulating layer 203 and the conductive layer 206 is about 200Å to 300Å.

In Fig. 3e and Fig. 4e, a source isolation layer 207 with a thickness between 500Å and 1000Å is then formed on the conductive layer 206, and limitation of an electrical connection between the source regions S is fully achieved by the conductive layer 206 therebelow. The source isolation layer 207 is formed by sequential deposition, etch-back, and recession of insulating materials such as silicon dioxide, for example. Next, a threshold voltage implantation V_t is performed on sidewalls of the long trench 250 to adjust the threshold voltage of each memory cell.

In Fig. 3f and Fig. 4f, a tunnel oxide layer 208 is formed on the two sidewalls of the long trench 250. The tunnel oxide layer 208 can be, for example, silicon dioxide formed by thermal oxidation with a thickness between 80Å and 150Å. Next, a conformable floating gate layer 209 is formed on the mask layer 202 and in the long trench 250 and contacts the tunnel oxide layers 208 thereby. Materials of the floating gate layer 209 can be, for example, N-type dopant doped polysilicon formed by LPCVD and the thickness thereof is between 300Å and 500Å.

In Fig. 3g and 4g, another protective layer 210 with a surface slightly below the surface of the substrate 200 is formed in the long trench 250 and exposes portions of the floating gate layer 209. The protective layer 210 is formed by sequential deposition, etch-back, and recession of materials such as boro-silicate-glass (BSG) formed by APCVD. The portions of the floating gate layer 209 exposed by the protective layer 210 are removed, leaving a U-shaped floating gate 209 in the long trench 250 as a floating gate and on the source isolation layer 207 and contacts the tunnel oxide layers 208. Then a conformable control gate

spacer layer 211 is formed on the mask layer 202 and in the long trench 250. The control gate spacer layer 211 can be, for example, silicon dioxide formed by LPCVD.

In Fig. 3h and Fig. 4h, the control gate spacer layer 211 is then etched, leaving control gate spaces 211 substantially having the same width as the vertical portions of the U-shaped floating gate 209 on sidewalls of the long trench 250. In addition, materials of the control gate spacer layer 211 and the protective layer 210 are similar, such as silicon dioxide and BSG here. Thus, during etching of the control gate spacer layer 211, the protective layer 210 is lightly etched and the thickness thereof is also decreased by the described etching.

In Fig. 3i and Fig. 4i, the protective layer 210 is then removed by, for example, wet etching, leaving control gate spacers 211. Next, a conformable inter-gate dielectric layer 212 formed by, for example LPCVD, is deposited on the mask layer 202 and in the long trench 250. The material of the inter-gate dielectric layer 212 can be, for example, silicon dioxide with thickness between 100Å and 200Å. Then a control gate 213 is formed in the long trench 250 and on the mask layer 202. The control gate 213 is formed by sequential deposition, etch-back, and recession of conductive materials such as N-type dopant doped polysilicon formed by LPCVD. The thickness of the deposited conductive materials is between 1000Å and 1500Å. Next, the control gate layer 213 and the inter-gate dielectric layer 212 are etched and a U-shaped inter-gate dielectric layer 212 and a control gate (the residual control gate layer 213) therein are formed in the long trench 250. The coupling ratio between the control gate and the floating gate can be

increased by the additional contact surfaces provided by the U-shaped inter-gate dielectric layer 209 therebetween.

In Fig. 3j and Fig. 4j, a plurality of parallel long isolation trenches are formed in the substrate 200 (shown in Fig. 4j) along a second direction, perpendicular to the first direction of the long trenches 250, by sequential lithography and etching. The described etching process stops at the source isolation layer 207 of the long trench 250 therein and a plurality of trenches 250' comprising the cell structures are thus defined in the substrate 200. This top view is shown in Fig. 2b. Fig. 3j shows a cross-section along the A-A' line where the trenches comprising cell structures are located and Fig. 4j shows a cross-section along the B-B' where follow-up shallow trench isolation (STI) regions are located.

In Fig. 3k and Fig. 4k, materials of an insulating layer 214 are filled in the described long isolation trenches. The insulating material can be, for example, silicon dioxide formed by high density plasma chemical vapor deposition (HDP CVD).

In Fig. 3l and Fig. 4l, a planarization process is then performed to level the wafer using the mask layer 202 as a polishing stop layer, leaving an insulating layer 214 therein and the shallow trench isolation (STI) region is thus formed. Next, the mask layer 202 is removed by, for example, wet etching, and the pad oxide layer 201 therebelow is exposed. Then a drain implantation (not shown) is performed on implant N-type impurities such as phosphorous (P) or arsenic (As) ions into the substrate 200. Then a thermal annealing process (not shown), for example a rapid thermal annealing (RTP) process, is performed and drain

regions D are thus respectively formed in the substrate 200 adjacent to each trench 250'. Then the pad oxide layer 201 is removed and a second insulating layer 215 is formed on each drain region D. The second insulating layer 215 is formed by sequential deposition and planarization of materials of the second insulating layer 215. After the planarization, the wafer surface is leveled and the control gate 213 is thus exposed.

In Fig. 1, a top view illustrating possible follow-up wordline BL and bitline WL is shown. In Fig. 1, the control gates 213 within the cell trenches (referring to trench 250') are further connected by the interconnecting hypothesis wordlines WL and hypothesis bitlines BL in phantom structures. In addition, the bitlines BL can also connect the drain regions D therebelow by a proper contact window (not shown) and the stacked gate flash memory devices are thus formed.

Compared with flash memory cell of the Prior Art, the present invention has the following advantages.

First, cells of the flash memory device in accordance with the invention are trench-type stacked gate flash memory devices disposed in cell trenches within a substrate rather than that normally disposed on the surface of a substrate in the Prior Art. Memory cell design of the invention can achieve higher integration of memory cell capacity than that in the Prior Art.

In addition, most of the fabricating processes in the invention are self-aligned. Thus, additional lithography processes and number of masks for the whole fabricating process can be reduced. The complexity of fabricating the

Stacked GATE flash memory device of the present invention is reduced and can be easily achieved.

Second, cells of the stacked gate flash memory device of the present invention are formed into the substrate.

5 Thus, the size of each flash memory cell can be minimized and the integration of the memory cell can be increased. Thus the capacity of a flash memory device can be increased and the current within a cell can also be increased by enlarging the depth of the cell trench. Furthermore, most
10 patterns of the masks for fabricating the stacked gate flash memory device are rectangular and can be easily fabricated. The costs of mask fabrication can be reduced and resolution limitations by the photolithography tools can be reduced.

While the invention has been described by way of
15 example and in terms of the preferred embodiments, it is to be understood that the invention is not limited to the disclosed embodiments. To the contrary, it is intended to cover various modifications and similar arrangements (as would be apparent to those skilled in the art). Therefore,
20 the scope of the appended claims should be accorded the broadest interpretation so as to encompass all such modifications and similar arrangements.